



Exploration & Seminars in AI (ABAI1009P).

# Amazon India 2025 Sales Performance & Pricing Strategy Analysis

Submitted to : Dr. Sakshi



## Team Members



**Anshika Bhaduria (A25ARIU0047)**



**Mohammad Suhan Khan  
(A25ARIU0049)**



**Siddhansh Jain  
(A25ARIU0050)**

# INTRODUCTION

## Background Story

In the hyper-competitive world of Amazon e-commerce, pricing is the lever that makes or breaks a business. Many sellers rely on guesswork, leading to lost revenue or unsold inventory. We recognized a need to bridge the gap between raw data and actionable strategy. Our goal was to empower sellers with a data-driven tool that reveals exactly how price fluctuations impact sales volume, moving away from intuition and toward precision..



## Technical Approach

- **Data Acquisition:** We sourced a comprehensive Amazon sales dataset from **Kaggle**.
- **Modeling:** We utilized **Google Colab** for data preprocessing and training a classification model to identify sales patterns relative to price points.
- **Deployment:** The source code was managed on **GitHub** and connected directly to **Streamlit Cloud**, rendering a live, interactive web interface for real-time predictions.

## Outcome

The result is a user-friendly Streamlit web application. Users can input specific pricing parameters, and the underlying model predicts the likely sales classification (e.g., High vs. Low volume). This tool visualizes the direct correlation between cost and demand, providing Amazon sellers with immediate insights to optimize their pricing strategies for maximum profitability.

# Problem Statement Objective

## Primary Objective

To quantify the correlation between product pricing and sales volume on Amazon India in 2025, enabling data-driven optimization of pricing models to maximize revenue.

- **Analyze Price Elasticity:** Determine how sensitive demand is to price changes for different product categories (e.g., identifying which goods are "elastic" vs. "inelastic").
- **Optimize for Profitability vs. Volume:** Distinguish between price points that drive maximum *sales volume* versus those that drive maximum *profit margin*.
- **Competitor Benchmarking:** (Optional but recommended) Compare internal pricing against competitor data to measure the "win rate" of the Buy Box at different price levels.







# Model Classification

Link to Model Classification



 [colab.research.google.com](https://colab.research.google.com)

**Google Colab**



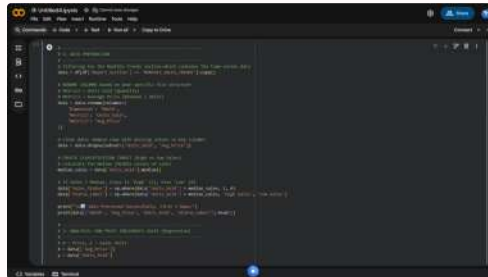
# Screenshots

1



```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import StandardScaler, MinMaxScaler, RobustScaler
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
```

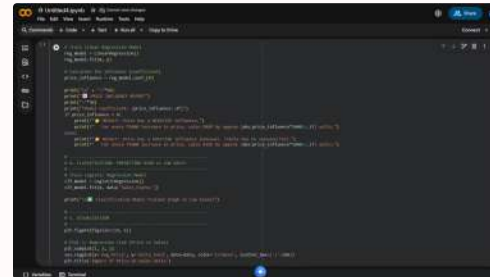
2



```
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

X_train, X_test, y_train, y_test = train_test_split(X_scaled, y,
                                                    test_size=0.2,
                                                    random_state=42)
```

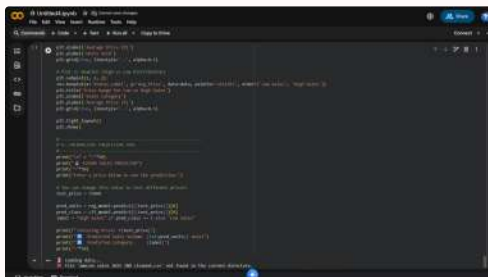
3



```
from sklearn.linear_model import LinearRegression

model = LinearRegression()
model.fit(X_train, y_train)
```

4



```
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
```

5



```
r2 = r2_score(y_test, y_pred)
```

6



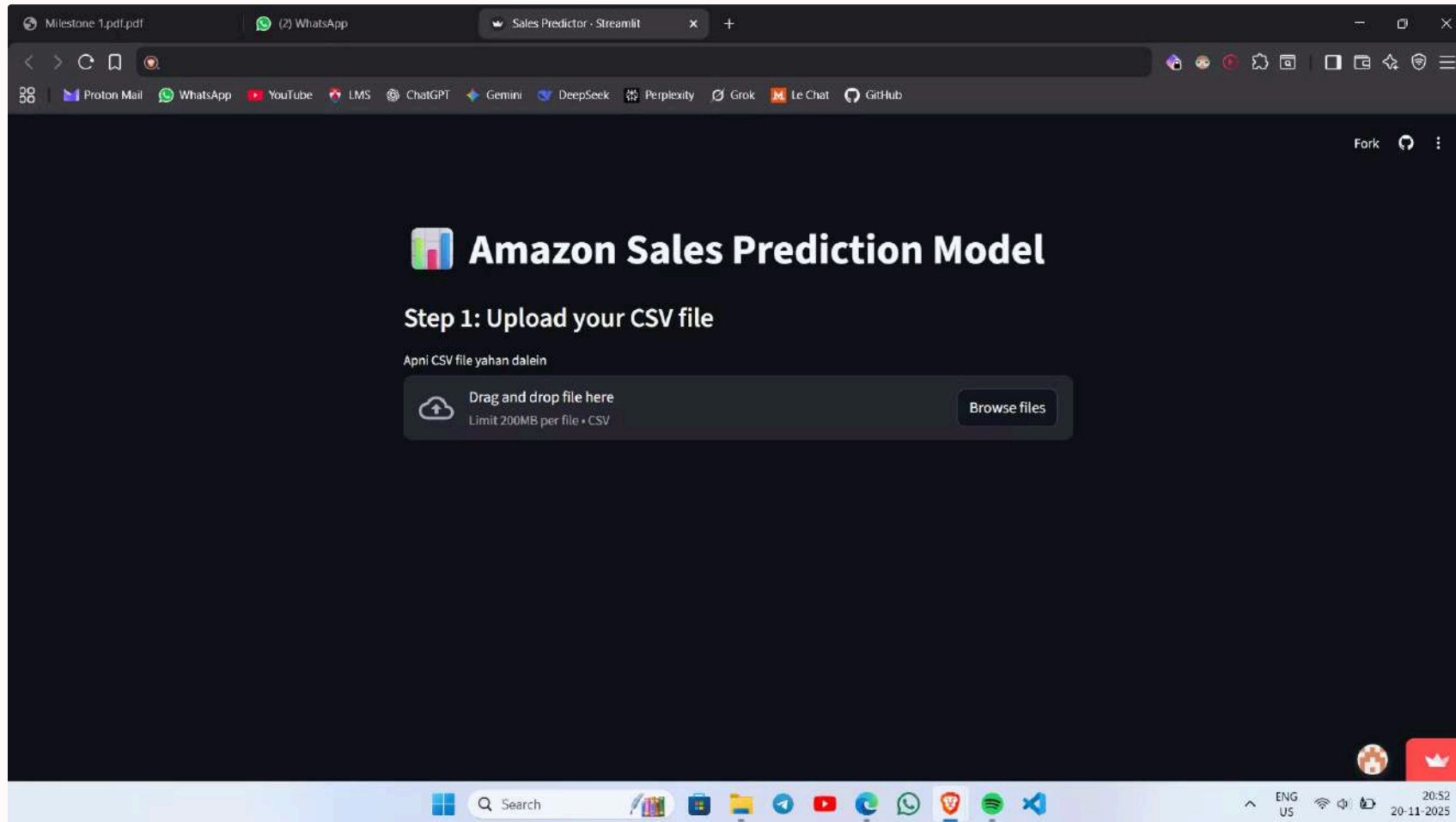


# Website Generation

**Link to the website :**

<https://amazonsalespredictor-hez79pp9ggpd9rdhqggnxr.streamlit.app/>

# Website Screenshot





# Amazon Sales Prediction Model: Analyzing Pricing Strategies in 2025

A Machine Learning approach to understanding Sales Elasticity on Amazon India

**Presenters:** Anshika Bhadauria, Mohammad Suhan Khan, Siddhansh Jain

**Context:** School of Artificial Intelligence | July-Dec 2025



# Abstract & Objectives: Unlocking Pricing Power

## Purpose

To analyze product sales performance on Amazon India and quantify the impact of pricing strategies on consumer purchasing decisions. Our goal is to empower sellers with actionable insights.

## Core Problem

Understanding pricing elasticity is paramount for vendor profitability in the rapidly maturing Indian e-commerce landscape. Suboptimal pricing directly impacts revenue and market share.



## Solution

An interactive, web-based tool (Streamlit) designed to analyze historical sales data and predict sales volume based on user-defined price inputs. This provides a dynamic simulation environment.



# Introduction: Navigating the E-commerce Wave in India



## Rapid Expansion

The e-commerce sector in India is experiencing exponential growth, presenting both immense opportunities and significant competitive pressures for sellers.



## The Pricing Conundrum

Sellers frequently grapple with identifying the optimal price point. Prices too high deter sales, while prices too low erode profit margins.



## Our Project Goal

We developed a robust system that ingests Amazon sales data, visualizes historical trends, and accurately predicts future sales performance using advanced machine learning models.

# Methodology & Workflow: Our Data Science Journey



## Data Collection

Sourced a comprehensive sales dataset representative of 2025 trends from Kaggle, ensuring rich historical context.



## Preprocessing

Executed rigorous data cleaning and feature engineering using Pandas within a Google Colab environment for efficient manipulation.



## Modeling

Implemented both Linear Regression and Classification algorithms to uncover nuanced patterns and relationships within the sales data.



## Deployment

Successfully hosted the final predictive solution as an intuitive web application using Streamlit, making it accessible to users.



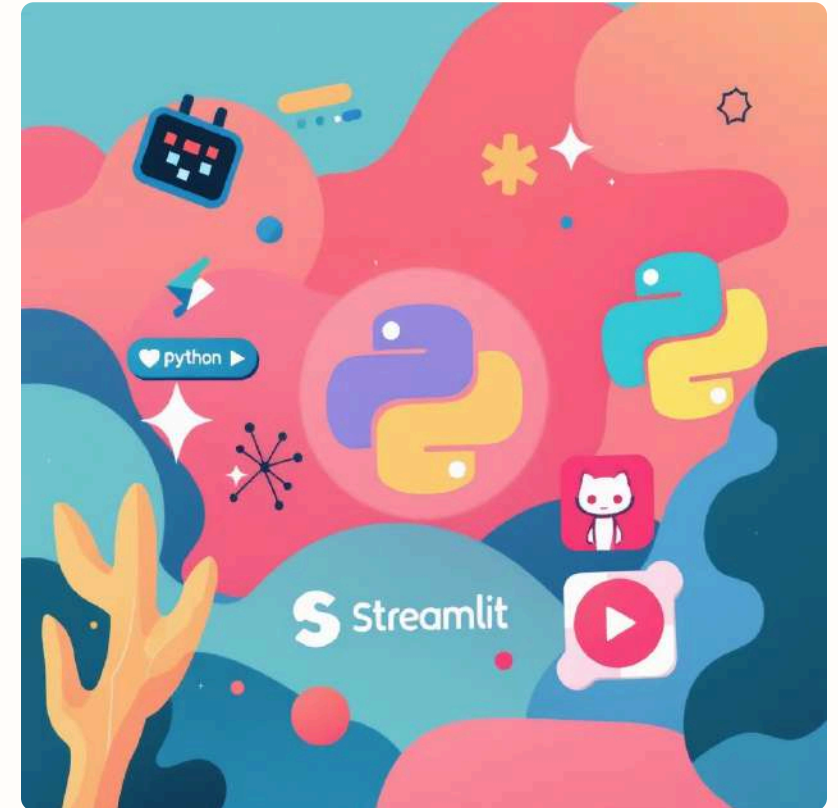
## Version Control

Managed all code collaboration, iterative development, and historical changes efficiently through GitHub, ensuring project integrity.



# Software & Technology Stack: The Tools of Our Trade

- **Python:** The foundational programming language for all development stages.
- **Google Colab:** Utilized for initial model training, extensive data filtering, and classification tasks.
- **Streamlit:** Empowered the creation of a user-friendly and interactive frontend interface for the web application.
- **Scikit-Learn:** Employed for implementing robust Linear Regression and Logistic Regression algorithms.
- **GitHub:** Served as the central repository for code hosting, version control, and collaborative development.
- **Libraries:** Pandas for sophisticated data manipulation; Matplotlib and Seaborn for compelling data visualization.







# Data Analysis & Visuals: Insights from the Amazon Marketplace

1

## Key Metrics Tracked

Our analysis focused on critical business metrics including Average Order Value (AOV), detailed Category Performance, and Customer Count segmented by State, providing a holistic view of sales dynamics.

2

## Regression Plot

A **Regplot** was employed to visually represent the linear relationship and trend between 'Average Price' and 'Units Sold', highlighting price elasticity.

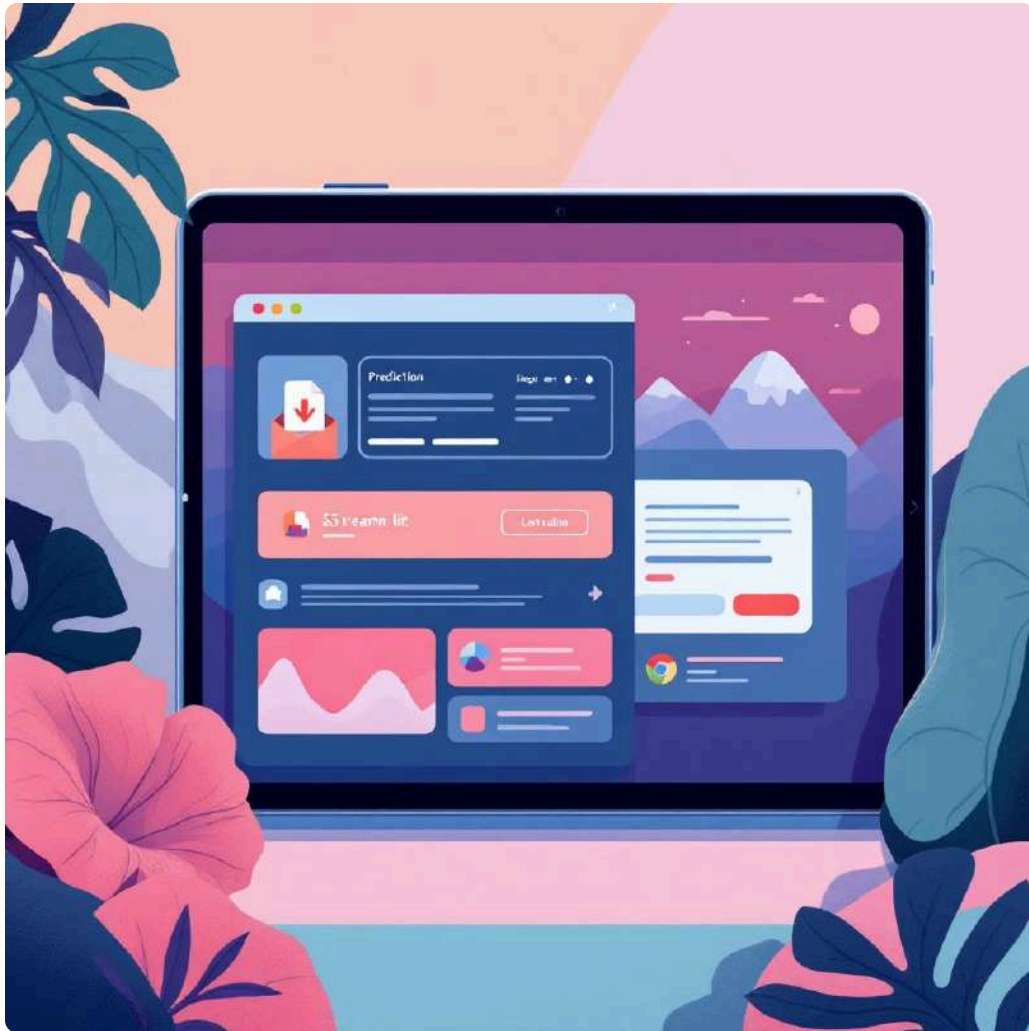
3

## Box Plots

**Boxplots** were instrumental in identifying distinct price ranges associated with 'High' versus 'Low' sales statuses, revealing optimal pricing tiers.

Through these visualizations, the system effectively identifies outliers in monthly sales data and discerns underlying general trends, enabling proactive adjustments.

# The Streamlit Application: Your Intuitive Pricing Command Center



## Interactive Features

- **File Uploader:** Seamlessly upload your proprietary CSV sales data for immediate analysis and prediction.
- **Metric Dashboard:** Get real-time insights into "Price Influence," quantifying how sales volume responds to every ₹1000 price alteration.
- **Future Predictor:** Utilize a dedicated input field to test "New Prices" and receive instant predictions of expected units sold, enabling proactive decision-making.

This application puts sophisticated data science directly into the hands of Amazon sellers, offering an unparalleled advantage in competitive pricing.

# Experimental Results: Validating Our Predictive Power

## Model Performance

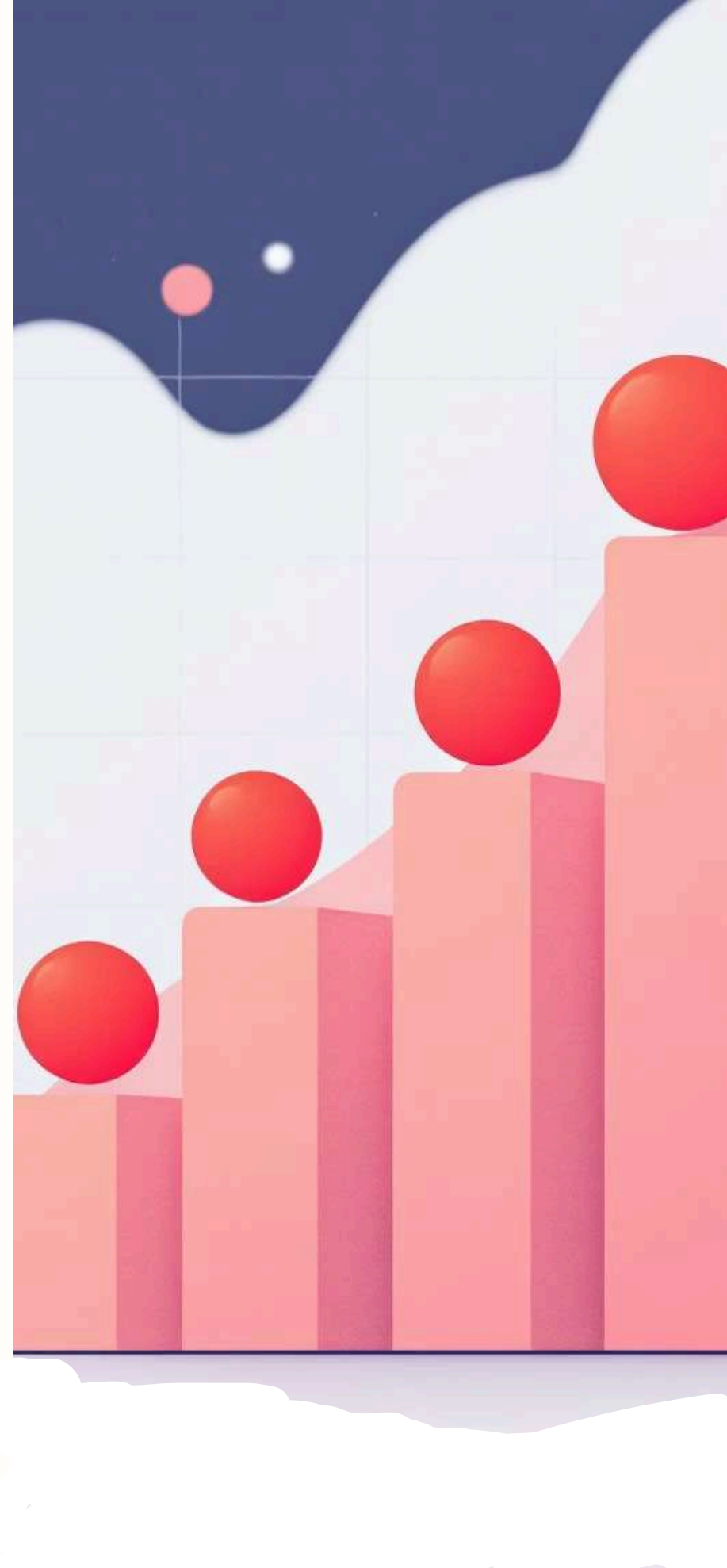
The Linear Regression model provided foundational insights, clearly indicating a negative correlation between product price and sales volume, as hypothesized.

## Key Finding: Beyond Price

While price is a significant determinant, our model revealed it's not the sole factor. Sales volume reacts distinctly and non-linearly to sharp price increases, suggesting nuanced consumer behavior.

## App Utility: "Predict Karo"

The "Predict Karo" feature has proven highly effective, allowing vendors to simulate diverse pricing scenarios and gauge potential sales outcomes **before** products are even listed. This translates to informed, strategic pricing.







# Conclusion: Empowering Amazon Sellers with Data Intelligence

## End-to-End Solution

We successfully engineered a complete pipeline, transforming raw Amazon sales data into actionable insights via a deployed web application, from ingestion to prediction.

## Democratizing Data Science

Our interactive tool makes advanced data science accessible to all Amazon sellers, enabling them to transition from speculative pricing to confident, data-driven decisions.

## Integrated Achievement

The project stands as a testament to successfully integrating complex data analysis, compelling visualization, and accurate predictive modeling into a single, intuitive interface.

# Future Scope: Evolving Our Predictive Capabilities



## User Authentication

Implement robust user authentication, allowing sellers to securely save and access their personalized prediction histories and preferences.



## Competitor Analysis

Develop a module to scrape and integrate real-time competitor pricing data, offering a comparative advantage to sellers.



## NLP Integration

Incorporate sentiment analysis of customer reviews to understand how perceived quality impacts sales in conjunction with pricing.



## Advanced Algorithms

Transition from Linear Regression to more sophisticated algorithms like Random Forest or XGBoost for enhanced predictive accuracy.